Heredity **82** (1999) 252–260 *Received 31 March 1998, accepted 8 September 1998*

Properties of AFLP markers in inheritance and genetic diversity studies of *Pinus sylvestris* L.

ESTELLE LERCETEAU* & ALFRED E. SZMIDT

Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-901 83, Umeå, Sweden

We analysed the properties of AFLP markers in *Pinus sylvestris*. Using primers with three selective nucleotides, the AFLP protocol produced large numbers of amplified bands and could only be used with a restricted number of primer combinations. Replacement of the *Eco*RI +3 primer by an *Eco*RI +4 primer halved the number of bands, facilitating analysis. The inheritance of all but about 8.4% of the amplified bands has been confirmed to be Mendelian. We compared band patterns among selected *P. sylvestris* trees from northern Sweden, two Asian species of *Pinus* and one species from the genus *Picea*. The dendrogram obtained was generally concordant with the taxonomic data, although the genetic similarity values between trees from different genera did not entirely follow accepted inter- and intraspecific relationships. This deviation was less pronounced using primer combinations that generated fewer bands. More than 69.1% of the bands that were polymorphic in two *P. sylvestris* trees or 29 of their F_1 progeny were in a pseudo-testcross configuration and thus were useful for the development of a linkage map for each parent. These markers have been analysed in four other crosses, and 83% of the bands could be mapped in at least one cross. Depending on the level of heterozygosity of the parents, the efficiency of such mapping will vary, but the AFLP technique appears to be a powerful way to generate, very quickly, large numbers of markers that are useful for constructing and comparing linkage maps.

Keywords: AFLP, conifers, genetic diversity, inheritance, mapping, *Pinus sylvestris* L.

Introduction

Pinus sylvestris L. (Scots pine) is among the most important timber tree species in many Eurasian countries. The main goals for Scots pine breeding in Sweden are the improvement of adaptation and survival, yield, and quality of the stem and wood.

In many plants, molecular markers have been found to be useful for studying genome organization, for evaluating the genetic structure of populations, for following evolutionary changes, and for assisting in traditional breeding. In various conifers, including *P. sylvestris* L., allozymes have been widely used in population genetics (e.g. Szmidt & Wang, 1993). However, attempts to develop linkage maps that would be useful in marker-assisted selection (e.g. Szmidt & Muona, 1989) have not been successful because the number of allozyme loci available is too small.

Devey *et al*. (1996) have developed linkage maps using RFLP markers in *Pinus radiata*. Because of the large size of the conifer genome and the high level of repetitive sequences, development of such maps is labour intensive and time consuming. Moreover, the inherent codominance and locus-specificity of RFLPs that make the method particularly attractive in many species are more of a hindrance in conifers because of the complicated patterns obtained (Devey *et al*., 1996).

Although the characterization of codominant and species-specific microsatellites is under way in many species, this development is slower and more difficult in conifers (Kostia *et al*., 1995). Too few microsatellites have been described per species to develop suitable linkage maps.

For genome mapping in conifers, RAPDs have been preferred and widely used (e.g. Yazdani *et al*., 1995). However, their reproducibility and stability across populations are often questioned (Heun & Helentjaris, 1993).

Because no marker is optimally suited for all studies, we have to develop a wide range of systems to be used either singly or in combination, depending on the requirement. Owing mainly to the large number of

*Correspondence. E-mail: estelle.lerceteau@genfys.slu.se

bands generated in a single reaction, the AFLP method (Vos *et al*., 1995) is believed to be a reliable and powerful tool for a wide variety of purposes including the evaluation of genetic variability (e.g. Roa *et al*., 1997) and the development of linkage maps (e.g. Waugh *et al*., 1997). In conifers, however, this method has been delayed because of the large size of the genome. The standard protocol has to be modified in order to decrease the number of bands generated per primer combination. Recently Paglia & Morgante (1998) reported the use of AFLPs in *Picea abies* using a strategy different from the one we developed.

In the study described below, we evaluated the AFLP method in *Pinus sylvestris*, by testing the inheritance of the amplified bands, and by checking whether variation of AFLP bands follows established inter- and intraspecific relationships. Implications of our findings for constructing and comparing genetic maps are also discussed.

Materials and methods

Plant material

The plant material analysed in this study is listed in Table 1. It included 10 *Pinus sylvestris* trees, originating in northern Sweden and currently used in the Swedish breeding programmes. Three progeny from the BD1032 × AC1014 cross and 29 from the AC3065 × Y3088 cross were also assessed, although the latter were only used for the evaluation of the segregation pattern of AFLP bands. This progeny constitutes part of the population which will be used to develop a genetic map of the species. The study also included a *P. merkusii* tree and a *P. gerardiana* tree, which belong to the *Pinus* (like *P. sylvestris*) and the *Strobus* subgenera of *Pinus*, respectively. A tree of *Picea abies* from the related genus *Picea* was also analysed.

AFLPs

Total DNA from needles or buds was extracted using a modified protocol (Doyle & Doyle, 1990) involving RNAase A (Boehringer) (final concentration 200 μg/mL) and Proteinase K (Merck) (final concentration 400 μg/mL).

The AFLP system (AFLP Analysis System I and AFLP Starter Primer Kit) from Life Technology AB was used. It is based on a two-step amplification strategy using *Eco*RI and *Mse*I primers (Vos *et al*., 1995). The bands were first amplified with primers each having one selective nucleotide. Then, the diluted PCR products

Table 1 Plant material used in the study

Name	Subgenus	Origin	Latitude	Longitude	Altitude (m)
Pinus sylvestris	*Pinus*				
AC1005		Northern Sweden	64°39′	10°06′	250
AC1014		Northern Sweden	64°53′	18°15′	330
AC1016		Northern Sweden	64°48′	18°39′	365
AC1019		Northern Sweden	64°14′	16°18′	320
AC3065		Northern Sweden	65°08′	20°14′	300
BD1032		Northern Sweden	66°28′	17°58′	445
BD1178		Northern Sweden	66°07′	20°40′	275
Y3001		Northern Sweden	63°48′	16°27′	230
Y3020†		Northern Sweden	64°13′	16°17′	275
Y3088		Northern Sweden	64°09′	16°04′	250
BD1032 × AC1014		Northern Sweden			
1801					
1802					
1804					
AC3065 × Y3088		Northern Sweden			
(29 progeny)					
Pinus merkusii	*Pinus*	Thailand			
Ms (13)					
Pinus gerardiana	*Strobus*	China			
IFG 28-2-0					
Picea abies		Sweden			
AC1033					

†Duplicated DNA.

were used as templates for the second amplification using primers containing three selective nucleotides (referred to as +3 primers). The procedure was performed as described in the kit, using γ-^{33}P ATP. We also tested the effect of adding another selective base (C) to the *Eco*RI primer (E-ACG), giving a '+4 primer', in the second amplification.

Following the amplification, reaction products were mixed with an equal volume of 98% formamide, 10 mM EDTA, bromophenol blue and xylene cyanol. After denaturation, 2 μL of each sample was loaded on a 6% denaturing (sequencing) polyacrylamide gel (Long Ranger gel, FMC). The gel was pre-equilibrated by passing an electric current through it (at a constant 70W) for 30 min. The reaction products were then separated, under these conditions, for 2 h 45 min. Following the separation, the gel was dried and exposed to X-ray film for 5–7 days. The presence and absence of the bands were then visually recorded.

Data analysis

Each marker analysed among the 29 F_1 progeny of the AC3065 × Y3088 cross was tested for the expected segregation ratio using a χ^2 goodness of fit test.

The material studied in the genetic divergence analysis included the 10 *P. sylvestris* trees, the three progeny from the BD1032 × AC1014 cross and trees from the related species and genus. For three primer combinations, (E-ACGC/M-CAC, E-ACG/M-CAG, E-AGC/M-CAC), the genetic similarity index (*GS*) of Nei & Li (1979) was computed. The dendrogram was established according to the genetic distance matrix ($D = 1 - GS$), following the unweighted pair group procedure with arithmetic mean (UPGMA) using the PHYLIP package version 3.5c (Felsenstein, 1993). The degree of relationship between the genetic similarity matrices generated with the different primer combinations was measured using the Mantel matrix-correspondence test (Mantel, 1967) implemented in NTSYS-PC version 1.60 (Rohlf, 1990). The test takes two matrices and plots one against the other and gives a product-moment correlation *r*. The significance of *r* is tested by a randomization procedure that compares the correlation between the two matrices with the correlation between one of these and the randomization of the other. The resulting *P* is the probability of H_o (no relationship) being right after 9999 permutations.

The *G*-test of independence was used as described by Sokal & Rohlf (1981) to test the independence between the segregation classes and the primer combinations.

Results

+3 primer screening and comparison of +3 and +4 primers

Eight *Eco*RI +3 primers (E-AAC, E-AAG, E-ACA, E-ACT, E-ACC, E-ACG, E-AGC, E-AGG) and eight *Mse*I +3 primers (M-CAA, M-CAC, M-CAG, M-CAT, M-CTA, M-CTC, M-CTG, M-CTT) are available in the kit. The whole set of 64 combinations was tested on the two *P. sylvestris* parental trees (AC3065 and Y3088).

The size of the amplified bands ranged from 50 to 550 bp. The number of bands varied from about 150 to more than 200, depending on the primer combination. For some combinations, the exact number of bands could not be given because some parts of the gel were smeared, especially for the high molecular weight bands. In general, the number of amplified products was smaller when there was a high C-G percentage in either of the primer sequences. Among the 64 primer combinations which could be tested, 12 generated easily readable patterns, 17 others could possibly be used, and 35 were excluded because of the complexity of the profiles they generated.

In order to get clearer patterns, the *Eco*RI +3 primer (E-ACG) was replaced with an *Eco*RI +4 primer (E-ACGC), while the *Mse*I +3 primer was retained. The addition of the extra selective nucleotide reduced the number of amplified bands by about 50% to 60–90 bands (see total band number column in Table 2) and increased band intensity. The patterns became clearer which made the reading easier. Few bands that were not amplified with the *Eco*RI +3 primer were detected with the *Eco*RI +4 primer

In order to check the reproducibility of the method, the analysis of one DNA sample (Y3020) was duplicated using the same DNA extraction techniques, in conjunction with two different digestion, ligation, preamplification and amplification protocols. The genetic similarity indices (*GS*) calculated between the two replications equalled 0.9857 and 1, respectively, for the two primer combinations tested (E-ACG/M-CAG and E-AGC/M-CAC), showing that the method was highly reproducible.

Segregation analysis

Inheritance of the AFLP markers was analysed in 29 progeny of the cross AC3065 × Y3088. Five primer combinations were evaluated (Table 2). In order to get a better reading of the high molecular weight bands when

Table 2 Segregation analysis on 29 progeny of the *Pinus slyvestris* cross (AC3065 × Y3088)

Parental genotypes (AC3065 × Y3088) Segregation ratio				*AA* × aa No	*aa* × *AA* No	*Aa* × *Aa* 3:1	*Aa* × aa 1:1	*aa* × *Aa* 1:1
Primer combinations	Total band number	Number of analysed bands†‡	Distorted segregation (%)	Bands Number†§	Bands Number†§	Bands Number†§	Bands Number†§	Bands Number†§
E-ACGC/M-CAC	63	27 (42.9)	7.4	1 (4.0)	0 (0)	6 (24.0)	8 (32.0)	10 (40.0)
E-AGCC/M-CTG normal run	67¶	24 (35.8)	4.2	2 (8.7)	1 (4.4)	5 (21.7)	5 (21.7)	10 (43.5)
E-AGC/M-CTG long run	201	24 (11.9)	8.3	1 (4.6)	3 (13.6)	5 (22.7)	4 (18.2)	9 (40.9)
E-ACG/M-CAC	193	57 (29.5)	12.3	1 (2.0)	2 (4.0)	11 (22.0)	15 (30.0)	21 (42.0)
E-ACG/M-CAG	166	55 (33.1)	9.1	2 (4.0)	2 (4.0)	9 (18.0)	12 (24.0)	25 (50.0)
E-AGC/M-CAC	153	39 (25.5)	5.1	2 (5.4)	0 (0)	11 (29.7)	10 (27.0)	14 (37.8)
Total	843	226 (26.8)	8.4	9 (4.4)	8 (3.9)	47 (22.7)	54 (26.1)	89 (43.0)

†In parentheses, results expressed in percentages.
‡Either polymorphic between the parents or segregating in the progeny.
§Fragments showing a distorted segregation ratio are not considered.
¶ Only the part not read on the long run is considered.

using three selective nucleotide primers, replicate samples amplified with the E-AGC/M-CTG primer combination were electrophoretically separated twice; once for the standard 2 h 45 min, and once for longer (4 h 25 min). Of the 25 bands analysed on the long-run gel, eight bands could have been read on the standard gel. Thus, with longer and better separation, 17 additional polymorphic loci among parents or progeny could be scored; a more than 3-fold improvement. Each band was tested for the expected segregation ratio using the χ^2-test. The proportion of bands which were either polymorphic between the parents or segregated in the progeny, and thus would be useful for mapping (referred to as number of analysed bands in Table 2), varied from 11.9 to 42.9%, with an average of 26.8%.

The following segregation ratios were expected, depending on the allelic state of the parents at the analysed locus: no segregation, 3:1 and 1:1 (Table 2). 'No segregation' would be expected in three situations where (i) both parents were dominant homozygotes ($AA \times AA$), (ii) one parent was a dominant homozygote and the other a heterozygote ($AA \times Aa$, $Aa \times AA$) and (iii) both parents were homozygotes for a different allele ($AA \times aa$ or $aa \times AA$). Cases (i) and (ii) could not be differentiated and are not included in Table 2 . The percentages of bands given for each configuration class in Table 2 refer to nondistorted bands that are either polymorphic between the parents or segregated in the progeny. Case (iii) represented on average 8.3% of the bands. The 3:1 segregation was obtained when both parents were heterozygotes ($Aa \times Aa$, intercross configuration). This was observed for 22.7% of the loci. The 1:1 segregation ratio, observed when one parent was heterozygous and the other homozygous ($Aa \times aa$ or $aa \times Aa$, pseudo-testcross configuration), was obtained for 69.1% of the loci. Of these, 26.1% resulted from segregation between female gametes ($Aa \times aa$ configurations) and 43% between male gametes ($aa \times Aa$ configurations). This difference was attributed to variation in the heterozygosity level of the two parents, Y3088 being more heterozygous than AC3065.

Considering the five classes of bands for which the genotypes of the parents could be defined (Table 2) , the heterozygosity level of AC3065 equalled 48.8%, and the heterozygosity of Y3088 was 65.7%. The heterozygosity level could be further evaluated using the 772 bands that did not show segregation distortion (Table 2). The following situations could not be differentiated: $Aa \times AA$, $AA \times Aa$ and $AA \times AA$. We assumed that the number of $Aa \times AA$ situations equalled the number of $Aa \times aa$ situations (54) and made the same assumptions for $AA \times Aa$ and $aa \times Aa$ (89). This gave heterozygosity level of 20.1% and 29.1%, respectively, for AC3065 and Y3088, the parents of the 29 progeny.

The *G*-test of independence revealed that the number of bands per segregation class was independent of the primer combination considered ($G = 12.238$, $P < 0.001$). The percentage of bands showing distorted segregation was on average 8.4% ($P < 0.05$), with results varying depending on the primer combination analysed (Table 2).

Genetic divergence analysis

Three primer combinations (E-ACG/M-CAG, E-AGC/M-CAC and E-ACGC/M-CAC) were used to test

whether the variation of the AFLP bands conformed to known intra- and interspecific relationships. These primer pairs generated 392 polymorphic bands: 158, 116 and 118 for E-ACG/M-CAG, E-AGC/M-CAC and E-ACGC/M-CAC, respectively. Among the 392 polymorphic bands, 24.2%, 6.1%, 15.6% and 17.6% were specific for *Pinus sylvestris*, *P. merkusii*, *P. gerardiana* and *Picea abies*, respectively. A total of 247 bands were amplified in *P. sylvestris*, of which 61.9% were polymorphic.

The whole set of bands was used to calculate genetic similarity (*GS*) indices. Within *P. sylvestris*, the *GS* values ranged from 0.7852 to 0.9873 and on average equalled 0.8582 ($\pm$0.0293). The most closely related trees were (i) AC1005 and Y3020, and (ii) the parent BD1032 and the three progeny of the cross BD1032 $\times$ AC1014. AC1005 and Y3020 could hardly be differentiated even with 392 bands. AC3065 and AC1019 were the most distantly related trees. The genetic similarity index between AC3065 and Y3088, the parents of the cross for which a map is being developed, equalled 0.8274 and was below the mean value for *P. sylvestris*.

Among species, *P. merkusii* was closer to *P. sylvestris* (*GS* varying from 0.4982 to 0.5540) and to *Picea abies* (*GS* = 0.1826) than to *Pinus gerardiana* (GS = 0.1732). The most distantly related taxa in the study were *Pinus gerardiana* and *Picea abies*. These results differed from the usual expectations of taxonomy, in which the genetic differences between species would generally be expected to be smaller than differences between genera. However, the general topology of the dendrogram based on 392 bands followed the classical taxonomy (Fig. 1). All the *P. sylvestris* trees were tightly associated and were clearly separated from the remaining species. Among *P. sylvestris*, no real cluster could be defined except one which associated the three progeny of the cross BD1032 $\times$ AC1014 and the parent BD1032.

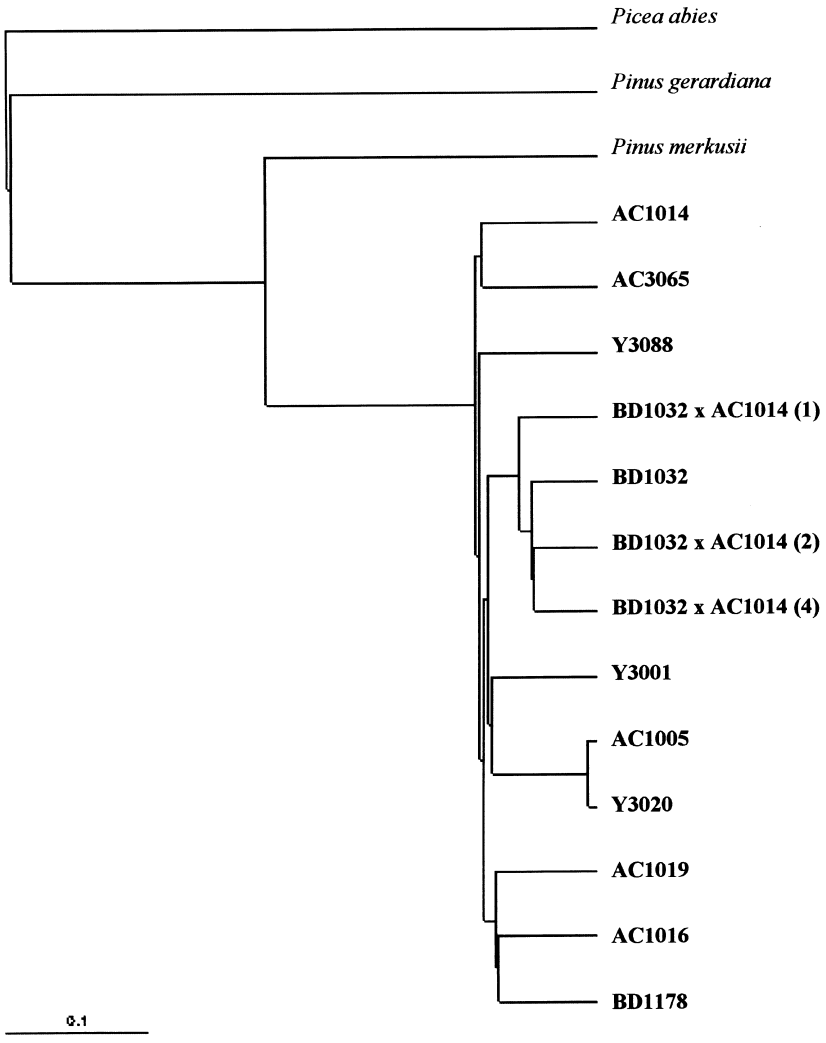

Fig. 1 UPGMA dendrogram of *Pinus sylvestris* accessions (bold), and *P. merkusii*, *P. gerardiana* and *Picea abies* (italic), based on (1−*GS*) genetic distance matrix with *GS* corresponding to the genetic similarity index of Nei & Li (1979).

The genetic similarity matrices generated from the three primer combinations individually considered or from the whole set of data (pooled results from the three primer combinations) were highly and significantly correlated (r ranging from 0.9834 to 0.9958, Mantel test $P < 0.0001$). Thus, the three primer combinations gave similar estimates of the genetic similarity among the trees tested. However, some relationships between some species differed according to the primer combination used in the analysis (Table 3); E-ACGC/M-CAC and E-AGC/M-CAC provided genetic similarities concordant with classical taxonomy, for instance. All of the primer combinations studied, however, showed the cluster composed of the parent BD1032 and its progeny, and a close relationship between AC1005 and Y3020. Thus, the strongest associations between trees were maintained, and did not depend on the primer pair.

Extrapolation to mapping other progeny

By comparing the amplification patterns obtained in the analysis of the 10 *P. sylvestris* trees, we checked how many bands mapped in one cross might be useful for mapping in other crosses. We chose randomly five crosses between the test trees: AC1014 × BD1032, AC3065 × Y3088,AC1019 × Y3001, AC1005 × AC1016 and BD1178 × Y3020. A total of 128 polymorphic bands were detected among these 10 trees; 50, 53, 47, 46 and 49 of which were polymorphic in the parents of the five crosses, respectively. Among the 128 polymorphic bands, 2.3% could be mapped in all five crosses, and 10.2%, 15.6%, 29.9%, and 31.3% could be mapped in four, three, two and one crosses, respectively. Among the 53 polymorphic bands observed from the AC3065 × Y3088 cross (for which a map will be developed), 5.7%, 18.9%, 24.5% and 34.0% could also be used to develop maps for four, three, two and one crosses of the other crosses, respectively. The remaining bands (17.0% of the 53) would be inappropriate for the crosses considered.

It should also be noted that a small percentage of the polymorphic bands corresponded to the configurations $AA \times aa$ or $aa \times AA$. As they will not reveal any polymorphism in the progeny, they will not be mapped. The frequency of $AA \times aa$ and $aa \times AA$ configurations will differ depending on the level of heterozygosity of the parents of the new crosses.

Discussion

Methodology improvement and reliability

We evaluated the properties of AFLPs in conifers in two ways. First, we used the standard protocol developed by Vos *et al.* (1995) for complex genome analysis (preamplification with +1 selective base primers and amplification with +3 primers), adjusting some parameters to improve separation of the bands. Secondly, we further modified the protocol by replacing the *Eco*RI +3 primer with an *Eco*RI +4 primer.

Using the first strategy, efforts were made to find the best migration conditions to obtain the highest resolution of the gel. The slightly modified protocol applied here proved to be successful, because 12 primer combinations out of 64 could be easily used. Furthermore, our protocol appeared to be highly reproducible when using +3 primer combinations on the duplicated sample ($GS > 0.98$). In *Picea abies*, Paglia & Morgante (1998) did not succeed with the same kit: they detected too many bands and very few discernible polymorphisms.

In the second strategy, we replaced the *Eco*RI +3 primer with an *Eco*RI +4 primer without changing the preamplification step (+1 primer), thus reducing the number of bands. Vos *et al.* (1995) showed the importance of the number of selective nucleotides in the reliability of the reaction. These authors found selectivity to be still acceptable with three (our case) selective nucleotides. In our experiments, the detection of some rare bands using the E-ACGC/M-CAC (+4 *Eco*RI) primer that were not detected using the E-ACG/M-CAC (+3 *Eco*RI) primer may reflect a certain lack of selectivity. This group of bands may also represent the less reproducible fraction of the amplified bands. The strategy developed tends to be supported first by a ratio of markers showing a distorted segregation with the +3/+4 pair (7.4%),

Table 3 Genetic similarity values at the interspecific and intergeneric levels for the three primer combinations analysed individually and on the whole date set

	Pinus merkusii/ *Pinus gerardiana*	*Pinus merkusii*/ *Picea abies*	*Pinus gerardiana*/ *Picea abies*
E-ACGC/M-CAC	0.1786	0.1212	0.0323
E-ACG/M-CAG	0.1237	0.2022	0.2222
E-AGC/M-CAC	0.2308	0.2093	0.0526
Whoe date set	0.1732	0.1826	0.1140

close to the total mean value (8.4%), and second by genetic similarity values generated at the interspecific level by the +3/+4 pair, which conform to classical taxonomy.

Both +3 and +4 *Eco*RI primers are informative. Even though 17 more primer pairs may have been potentially useful in the first strategy, the use of +4 *Eco*RI primers may increase if more primer combinations have to be tested as in a mapping project. In order to reduce the band number and despite the problems linked with methylation, Paglia & Morgante (1998) chose to replace *Eco*RI by *Pst*I which does not cut the repetitive sequences, thereby keeping three selective nucleotides.

We also tested the Mendelian inheritance of AFLPs derived using both the strategies we tested, because it is a prerequisite for their utilization in any kind of inheritance study. Segregation departure from Mendelian ratios can be caused by methodological problems, may have a biological meaning (e.g. difference in the viability of gametes carrying different alleles) or can occur by chance, as shown by Grattapaglia & Sederoff (1994). The number of distorted loci we found (8.4%) was intermediate between the frequency reported for *P. sylvestris* by Szmidt & Muona (1989) with allozymes (10.6%) and Yazdani *et al.* (1995) with RAPDs (5%), and thus validates the method. Maheswaran *et al.* (1997), and references therein, did not observe any difference in the distortion ratios detected using AFLP and RFLP markers to analyse the same population of rice.

Properties of AFLPs in genetic divergence studies

The high association among the genetic distance matrices generated using different primer combinations ($r > 0.98$) showed that our results were coherent and that a single pair of primers should be enough to obtain precise information.

The general topology of the final dendrogram followed classical taxonomy; *P. sylvestris* was separate from the other species. The lower taxonomic levels were also maintained, as *P. sylvestris* was found to be genetically closer to *P. merkusii* (subgenus *Pinus*) than to *P. gerardiana* (subgenus *Strobus*). Based on the analysis of the chloroplast (Cp) DNA of 24 *Pinus* species, Wang & Szmidt (1993) found similar results.

When considering a higher taxonomic level, it appeared that *Pinus merkusii* and *Pinus sylvestris* were closer to *Picea abies* than to *Pinus gerardiana*. One has to consider the method itself to explain this result. It is generally accepted that the AFLP product sizes and the selective part of the primer are reliable indicators of homology at least within a species (Waugh *et al.*, 1997). However, interspecific comparisons have generally shown a higher level of polymorphism, associated with an increased probability of obtaining size matches between nonhomologous products (Waugh *et al.*, 1997). In *Pinus*, a large number of bands were generated, making size comparison more difficult and thus the estimates of relatedness probably less reliable.

At the intraspecific level, the relationships between trees may be restored to those established by classical taxonomy. With the exception of the association between BD1032 and the three progeny of the cross between BD1032 and AC1014, no cluster was observed. Based only on material from northern Sweden, Szmidt & Wang (1993) and Szmidt *et al.* (1996) reported G_{ST} values of 0.004 and 0.022. The population differentiation in northern Sweden is probably too low and the intrapopulation variability too high to generate a visible cluster in the present study.

As AFLPs can quickly generate a large number of markers, much of the genome can be analysed. In this way, results may be more representative of the genome as a whole. However, because of their dominance, AFLPs like RAPDs will probably not fully reflect genetic relationships. Szmidt *et al.* (1996) showed in *P. sylvestris* that if the population is not in a Hardy–Weinberg equilibrium, a strong bias is observed when using RAPDs instead of isozymes to calculate the population-genetic parameters of diploid tissues. In conifers, utilization of the haploid megagametophytes can circumvent this difficulty. By analysing this tissue in different seeds from the same tree, the genotype of the parental tree can be deduced. It makes AFLPs more attractive for that purpose but also increases the cost of the experiment.

AFLPs in genetic mapping

For conifers, most maps previously developed rely on the use of the haploid megagametophytes. However, because they represent a temporary phase, their utilization in QTL mapping is limited to juvenile families established especially for that purpose (Plomion *et al.*, 1996). A good alternative is to use a pseudo-testcross approach on a full-sib population combined with dominant markers, as Grattapaglia & Sederoff (1994) did with RAPDs. The efficiency of the method depends on the number of loci in a testcross configuration ($Aa \times aa$), which is a function of the divergence between the parents (Grattapaglia & Sederoff, 1994). Our preliminary study of 29 progeny showed that 69.1% of the analysed and nondistorted AFLPs, and 75.2% of the segregating bands, were present in this configuration. For the cross we considered, the paternal genome would be easier to map than the maternal one. The percentage of loci showing a 3:1 segregation (intercross bands) was

nearly as high as the fraction of maternal testcross configurations. Markers heterozygous in both parents could be located on both maps, and thus may be useful for aligning the two parental maps. Unfortunately, the information content gained from pairs of markers segregating in 1:1 and 3:1 ratios is low (Ritter *et al*., 1990). Because only the homozygous recessive genotypes are used in the analysis, only one-quarter of the population data is useful for determining the positions of the markers heterozygous in both parents. Consequently, the location of these markers can be defined much less precisely than markers heterozygous in only one parent. Verhaegen & Plomion (1996) were able to map only 25% of the intercross bands detected between *Eucalyptus urophylla* and *E. grandis* for this reason. Despite the high number of intercross configurations observed, the cross chosen to develop a map of *P. sylvestris* proved to be very suitable, as the *GS* index calculated between the two parents was below the average for all 13 trees.

The efficiency of the pseudo-testcross strategy with AFLPs should be better than with RAPDs, because of the high number of bands that can be amplified with one primer combination, saving time and money. Thus, AFLPs may prove to be the method of choice for generating genetic maps (Mackill *et al*., 1996).

An important point, however, is to determine how useful information collected from one cross may be for analysing another cross. AFLPs are locus-specific and stable across populations (Waugh *et al*., 1997). Thus, the ability to align AFLP maps will depend on the number of alleles shared among the loci segregating in different populations. Because AFLPs generally act as dominant markers, for each locus only two alleles can be scored. The number of alleles per allozyme locus in some populations studied from northern Sweden is on average 3.2 (Szmidt & Wang, 1993). Based on these figures, it is obvious that by using dominant instead of codominant markers, some of the polymorphism will be undetected. However, our estimates show that 83% of the polymorphic AFLP bands between AC3065 and Y3088 could be mapped in at least one population among five derived from other crosses chosen at random. It should therefore be possible to identify linkage groups sharing common markers and to align maps from different crosses as has already been done in barley (Waugh *et al*., 1997). According to these authors, this circumvents the need of other types of single-locus (codominant) markers for chromosomal alignment. In our case, the basic strategy implies the use of AFLPs with populations in a two-way pseudo-testcross configuration. Even if some markers heterozygous for both parents can be added, the best way to make bridges between the intraparental maps may be to add codominant markers. Only later, by accumulating data on different crosses, will it be possible to associate the linkage groups derived from the parental maps using AFLPs alone.

Acknowledgements

We thank Maria M. Ribeiro for her help in the statistical analysis and Dr Bengt Andersson from the Forestry Research Institute of Sweden for helping choose a part of the analysed material. This study was sponsored by the Centre for Forest Biotechnology and Chemistry (Sweden) and by the Swedish Council for Forestry and Agricultural Research (SJFR).

References

DEVEY, M. E., BELL, J. C., SMITH, D. N., NEALE, D. B. AND MORAN, G. F. 1996. A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. *Theor. Appl. Genet*., **92**, 673–679.

DOYLE, J. J. AND DOYLE, J. L. 1990. Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.

FELSENSTEIN, J. 1993. PHYLIP (*Phylogeny inference package*) version 3.5c. Department of Genetics, University of Washington, Seattle. Distributed by the author.

GRATTAPAGLIA, D. AND SEDEROFF, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics*, **137**, 1121–1137.

HEUN, M. AND HELENTJARIS, T. 1993. Inheritance of RAPDs in F_1 hybrids of corn. *Theor. Appl. Genet*., **85**, 961–968.

KOSTIA, S., VARVIO, S. L., VAKKARI, P. AND PULKKINEN, P. 1995. Microsatellite sequences in *Pinus sylvestris*. *Genome*, **38**, 1244–1248.

MACKILL, D. J., ZHANG, Z., REDONA, E. D. AND COLOWIT, P. M. 1996. Level of polymorphism and genetic mapping of AFLP markers in rice. *Genome*, **39**, 969–977.

MAHESWARAN, M., SUBUDHI, P. K., NANDI, S., XU, J. C., PARCO, A., YANG, D. C. AND HUANG, N. 1997. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *Theor. Appl. Genet*., **94**, 39–45.

MANTEL, N. A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res*., **27**, 209–220.

NEI, M. AND LI, W.-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A*., **76**, 5269–5273.

PAGLIA, G. AND MORGANTE, M. 1998. PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. *Mol. Breed*., **4**, 173–177.

PLOMION, C., DUREL, C. E. AND VERHAEGEN, D. 1996. Marker-assisted selection in forest tree breeding programs as illustrated by two examples: Maritime pine and eucalyptus. *Ann. Sci. For*., **53**, 819–848.

RITTER, E., GEBHARDT, C. AND SALAMINI, F. 1990. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. *Genetics*, **125**, 645–654.

ROA, A. C., MAYA, M. M., DUQUE, M. C., TOHME, J., ALLEM, A. C. AND BONIERBALE, M. W. 1997. AFLP analysis of relationships among cassava and other *Manihot* species. *Theor. Appl. Genet.*, **95**, 741–750.

ROHLF, F. J. 1990. NTSYS-pc. *Numerical taxonomy and multivariate analysis system*. 1.60 edn. Exeter Software, Setauket, NY.

SOKAL, R. R. AND ROHLF, F. J. 1981. *Biometry*. W.H. Freeman and Co., San Francisco.

SZMIDT, A. E. AND MUONA, O. 1989. Linkage relationships in Scots pine (*Pinus sylvestris* L.). *Hereditas*, **111**, 91–97.

SZMIDT, A. E. AND WANG, X.-R. 1993. Molecular systematics and genetic differentiation of *Pinus sylvestris* (L.) and *P. densiflora* (Sieb. et Zucc.). *Theor. Appl. Genet.*, **86**, 159–165.

SZMIDT, A. E., WANG, X.-R. AND LU, M.-Z. 1996. Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L) using haploid tissue analysis. *Heredity*, **76**, 412–420.

VERHAEGEN, D. AND PLOMION, C. 1996. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. *Genome*, **39**, 1051–1061.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VANDELEE, T., HORNES, M. *ET AL.* 1995. AFLP: A new technique for DNA fingerprinting. *Nucl. Acids Res.*, **23**, 4407–4414.

WANG, X. R. AND SZMIDT, A. E. 1993. Chloroplast DNA-based phylogeny of Asian *Pinus* species (*Pinaceae*). *Pl. Syst. Evol.*, **188**, 197–211.

WAUGH, R., BONAR, N., BAIRD, E., THOMAS, B., GRANER, A., HAYES, P. AND POWELL, W. 1997. Homology of AFLP products in three mapping populations of barley. *Mol. Gen. Genet.*, **255**, 311–321.

YAZDANI, R., YEH, F. C. AND RIMSHA, J. 1995. Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. *For. Genet.*, **2**, 109–116.